# An Evaluation of the Forecasting Performance of ARIMA Models for Seasonally Adjusted and Unadjusted Data

**Merdi Ahmed Orsud**

Deparment of Statistics, Faculty of Mathematical and Computer Sciences, University of Gezira

## ABSTRACT

This paper undertakes an evalution of the forecasting performance of univariate ARIMA model when the data are seasonally adjusted and when they are not. The method of adjustment is the U. S. Bureau of Census Method X-II. To attained stationarity, natural logarithm transformation for water flow amount at Eldaim Station was taken before univariate ARIMA models for the adjusted and unadjusted data constructed (during 1990 to 2006) and their forecasting performance upon the process of updating was compared (during 2001 to 2006). The paired t-test for the actual and predicted monthly averages shows no difference between the univariate ARIMA models for the unadjusted and the adjusted data. The comparison of the forecast error statistics obtained reveals that the forecasting performance of the models for the unadjusted logged series is better than that of the adjusted. Therefore, the use of unadjusted monthly data when constructing ARIMA models for forecasting or control of water flow amounts at Eldaim Station is better.

**Key Words:** ARIMA, Adjusted, Unadjusted, Bureau of Census Method X-II.

## INTRODUCTION

When constructing univariate time series models, as those of Box and Jenkins (1976), forecasters can choose either a seasonal autoregressive integrated moving average (ARIMA) model based on the seasonally unadjusted data, or a nonseasonal ARIMA model based on the adjusted data.

Generally, model builders who use seasonally adjusted data wish to perceive the short-term trend and cyclical movements of the series in question undistorted by the more or less regular month to month variation. They assume that the seasonal effects of the data are best captured by the seasonal adjustment procedures than by an explicit seasonal ARIMA representation. Thus, users of seasonally adjusted data expect to obtain more reliable forecasts than those obtained from the unadjusted data( Musa, 1998).

Harvey (1981) argued that working with seasonally adjusted data is not generally desirable. This is because there is always the fear of over adjustment. Wallis (1974) concluded that seasonal adjustment can introduce considerable distortion into series, and, in addition, there is no guarantee that the seasonal effects will be entirely removed. Plosser (1979) also showed that within the class of ARIMA models, the use of seasonal adjusted data does not yield consistent improvement in forecast accuracy relative to forecasts based on the unadjusted data.

This paper was, therefore, deal with the problem of choosing between the seasonally adjusted and unadjusted data when constructing univariate time series models. And the objective of this paper is to test for the claimed improvement of seasonal adjustment in the performance of forecasting modes. The models are to be constructed via studying the history and the variability pattern of the data for case when it is seasonally adjusted and unadjusted. The paper expected that, working with seasonally adjusted data has no benefit over working with the unadjusted data, regarding the forecasting performance of univariate ARIMA models.

The data series is analyzed in order to compare the forecasts generated from models based on both the unadjusted and the seasonal adjusted forms of the data. For each form the data are utilized for the period Jan 1990 through Dec 2006. Adjustments to the outlying data points and transformations for attaining homoscedasticity are initially performed in the data series. The Box and Jenkins iterative techniques using data from Jan 1990 through Dec 2000 are then applied to the resulting series to find univariate time series models, firstly, for the original form, and secondly, for the adjusted from that results after applying the U. S. Bureau of the Census program. These models are then used to generate forecast of 2001 (i. e. Up to twelve month ahead). The models (in their initial form) are then reestimated with data from 1990 through 2001 and used to forecast the series for 2002. This process of making yearly forecasts and updating the parameter estimates is repeated for six years, trough 2006. The twelve monthly forecasts for each year are converted to an annual forecast representing the average of the monthly values predicted for the year. The annual forecasts obtained from the adjusted and the unadjusted data for the years 2001 through 2006 are then compared to their corresponding actual values by applying a pair t-test. The pair t-test is also applied to compare the monthly forecasts of the adjusted and unadjusted data for the period 2001 through 2006. The forecast error statistics obtained from the adjusted and the unadjusted data through the process of updating are compared to evaluate the relative performance of the predictions derived from the two forms of data. The comparison is in terms of forecast error means and standard deviations, mean absolute error and root mean squared error.

In the following we give some information about nonseasonal and seasonal Autoregressive Integrated Moving Average (*ARIMA*), seasonal adjustment and U. S. Bureau of the Census Program.

**Autoregressive Integrated Moving Average (*ARIMA*) Model**

In many practical situations the assumption of stationarity of time series is too restrictive, in that their plots will frequently show some kind of trend in the mean and possibly in the variance (Harvey, 1981).

Fortunately, most of the non-stationary series encountered can transform into stationary series. Such series are termed homogeneous non-stationary (Pindyck & Rubinfeled, 1981). Homogeneous non-stationary series transformed into stationary by successive differencing (Box & Jenkins, 1976); i.e., by considering $\nabla X_t$ , $\nabla^2 X_t$ , where $\nabla = 1- B$ is the difference operator. A generalization of the Autoregressive Moving Average (*ARMA*) models incorporating the above type of non-stationarity is given by the class of autoregressive integrated moving average (*ARIMA*) process and is defined as follows

If d is a non-negative integer, then $X_t$ is said to be an *ARIMA(p, d, q)* process if

$$\phi(B)\ \nabla^d X_t = \theta(B)\ a_t; \qquad (1)$$

Where $\phi(.)$ and $\theta(.)$ are stationary *AR* and invertible *MA* polynomials of degree *p* and *q* respectively.

Seasonality is observed from the time plot, which show some degrees of correlation with the corresponding data point which lead by s, and autocorrelation function, which will exhibit peaks at corresponding correlated data points.

The pure seasonal autoregressive moving average model, $ARMA(P, Q)_s$, takes the form

$$\Phi_P(B^s)\ X_t = \Theta_Q(B^s) a_t \qquad (2)$$

where

$$\Phi_P(B^s) = 1 - \Phi_1(B^s) - \Phi_2(B^{2s}) - \dots - \Phi_P(B^{Ps})$$

$$\Theta_Q(B^s) = 1 - \Theta_1(B^s) - \Theta_2(B^{2s}) - \dots - \Theta_Q(B^{Qs})$$

However, time series observations usually involve pattern other the seasonal movements. Multiplicative seasonal ARMA process incorporate such other pattern or relations denoted by $ARMA(p, q)\times(P, Q)_s$ and written as

$$\Phi_P(B^s)\ \phi(B) X_t = \Theta_Q(B^s)\theta(B) a_t \qquad (3)$$

where $\phi(B)$ and $\theta(B)$ are the ordinary autoregressive and moving average operators respectively (Shumway and Stoffer, 2000).

Since in seasonal data, the observations which are $s$ intervals apart are similar, a difference of order $s$ to the observations $X_t, X_{t-1}, X_{t-2}, \ldots,$ etc., will be of a particular importance in removing non-stationarity that may be found in the seasonal series. Using the backward shift operator, $B^s X_t = X_{t-s}$, a seasonal difference of order $D$ is denoted by $\nabla^D X_t = (1-B^s)^D X_t$. It may also happen that besides the seasonal differencing, a series may require a conventional differencing of order d, $\nabla^d$, to attain stationarity. Applying the two differencing operators to a multiplicative seasonal *ARMA* process, results in what Box and Jenkins (1976) call as a multiplicative seasonal *ARIMA* process of order $(p, d, q)\times(P, D, Q)_s$, and is written as

$$\Phi_P(B^s)\,\phi(B)\,\nabla^d\,\nabla^D\,X_t = \Theta_Q(B^s)\theta(B)a_t \qquad (4)$$

Box and Jenkins methodology is used for identifying, estimating and diagnosing seasonal and non-seasonal *ARIMA* models.

**Seasonal Adjustment**

The seasonal adjustment procedures are based on the idea that a time series $y_t$ can be represented as the product of components.

$$X_t = T \times S \times C \times I \qquad (5)$$

where

$T \equiv$ value of the long-run secular trend in series,

$S \equiv$ value of seasonal component,

$C \equiv$ value of cyclical component,

and $I \equiv$ irregular component.

Seasonal adjustment procedures are techniques used for computing the seasonal indices that measure the seasonal variation in the series. The indices are then used to deseasonalize or seasonally adjust the series by eliminating those seasonal variations (Pindyck and Rubinfeld, 1981). The common reason to remove seasonal components to leave a series with simpler pattern to be studied for its implications. Another reason is to make comparisons of series with different seasonal pattern (Cleveland and Tiao, 1976).

To eliminate the seasonal component $S$, the combined long-term trend and cyclical components $T \times C$ are first isolated by applying a smoothing filter that will approximately remove the combined seasonal and irregular components $S \times I$ form the original series $X_t$. Now that, the combined seasonal and irregular components can be obtained by dividing the original series by dividing the original series by the resulting smoothed $T \times C$ estimates, That is,

$$(T \times S \times C \times I)/(T \times C) = S \times I \qquad (6)$$

To eliminate the irregular component I in order to obtain the seasonal indices, the values $S \times I$ corresponding to the same month over the whole series are averaged.

The seasonally adjusted data of the original series $X_t$, are now obtained by dividing each of $X_t$ values by its corresponding seasonal index.

The seasonal adjustment procedures approximately follow the same logic of the description given above although each has its specific smoothing filter; whether symmetric or asymmetric.

For a series $X_1, X_2, \ldots, X_T$, of the length $T$, which is to deseasonalized, the resulting series $Y_t$ for $m+1 \le t \le T-m$ is obtained by

$$\mathrm{Y_t} = a_m(B)X_t = \sum_{j=-m}^{m} a_{m,j}X_{t-j} \qquad (7)$$

where $B$ is the lag operator, $a_m(B)$ is a linear filter equivalent to 2m+1-term moving average, and $a_{m,j} = a_{m,-j}$.

For the current and recent data (T-m ≤ t ≤ T) which cannot be computed by the filter $a_m(B)$, truncated, asymmetric filter $a_i(m)$ must be applie$\mathrm{Y_T^{(0)}} = a_0(B)X_T = \sum_{j=0}^{m} a_{0,j}X_{T-j}$ (8)

$$\mathrm{Y}_{\mathrm{T-m}}^{(\mathrm{m})} = a_m(B)X_{T-m} = \sum_{j=-m}^{m} a_{i,j}X_{T-m-j} \quad .$$

For the filter $a_i(B)$, $i$ = 0, 1, 2, ..., m, the subscript $i$ indicates the number of future values of $X$ entering the moving average.

One of the most widely used of the adjustment procedures is the Census II method developed programme X-11 which was developed by the bureau of the Census of the U. S. Department of Commerce (Shiskin et. al. 1967; McKenzie, 1984). It assumes the additive decomposition

$$X_t = P_t + S_t + e_t \qquad (9)$$

where $X_t$ is the observed series, $P_t$ is the trend component, $S_t$ is the seasonal component, and $e_t$ is a white noise.

$$\mathrm{Y}_{\mathrm{t}} = a_m(B)X_t = \sum_{j=-m}^{m} a_{m,j}X_{t-j}$$

The X-11 program computes the estimated values for $P_t$, $S_t$, and $e_t$, as $\bar{P}_t, \bar{S}_t,$ and $\bar{e}_t$ respectively. The multiplicative version of the program is obtained by applying an additive model to the logarithms of the series $X_t$. The program basically uses symmetric moving average filters with weights summing to one. A deseasonalized series, $Y_t$, for $t$ sufficiently far removed from the end of the series $X_t$, can thus be obtained as (Musa, 1998)

where $B$ is the lag operator, $a_m(B)$ is the linear filter equivalent to 2m+1 moving average, and $a_{m,j}$ are the weights used in averaging $X_t$, such that $a_{m,j}= a_{m,-j}$.

For the values at the ends of the series, the program uses asymmetric filters as described in equation (8).

The Census program uses 9-term, 13-term, or 23-term filters, and 13-term filter is the most used because it seems to be the one applicable to the majority of series.

**Models Building**

**An ARIMA Model for the Unadjusted Data**

Model Identification

Before building an ARIMA model for time series data, it is necessary to inspect the original series plot. This will help making adjustments and selecting the appropriate transformations for the series (Mabert, 1974).

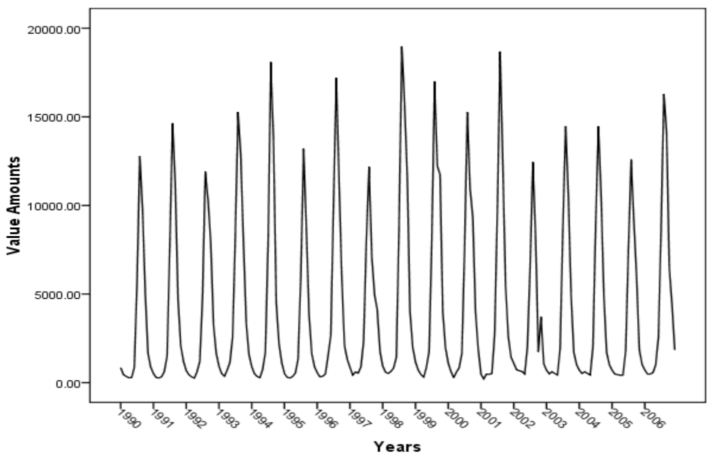


Fig. 1: The Plot of the Monthly Water Flow Amounts at Eldaim Station

The plot of the original series (Fig. 1) shows a marked seasonal pattern since the water flow is at its highest in autumn months and lowest in summer months. The yearly peaks of the autumn months exhibit a rather systematic decreasing and increasing trend, indicating that the series is not homoscedastic. To insure homoscedastic a logarithmic transformation applied to the original data. The plot of the logged series given in Fig. (2) Shows that the logged series is homoscedastic with approximately constant mean and, therefore, doesn't need a transformation.

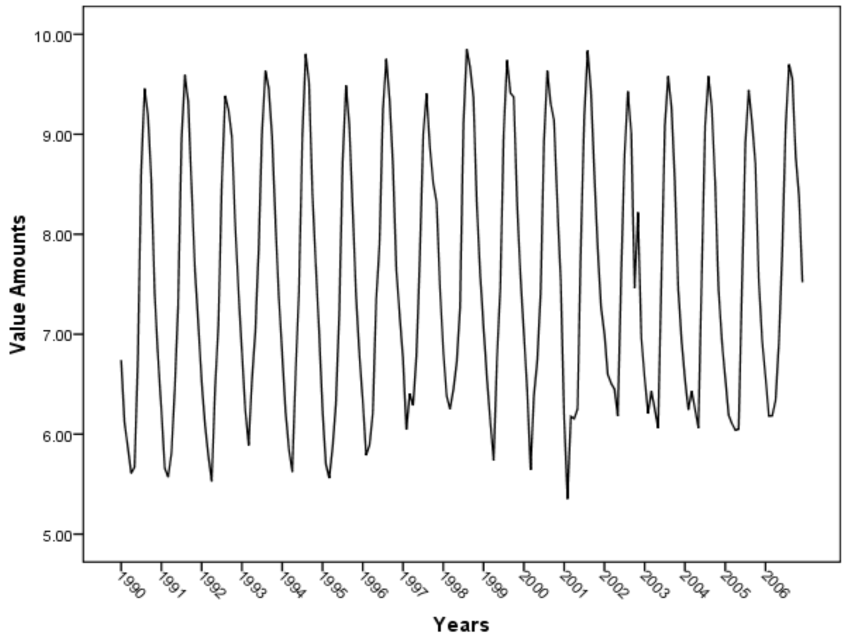



Fig. 2: Natural Logarithm of Monthly Water Flow at Eldaim Station

The logged series extending from January 1990 to December 2006 is divided into two parts. Part one, from January 1990 to December 2000, will be used to identify, estimate, and diagnostically check the tentative models suggested during the identification stage. The second part, from January 2001 to December 2006, will be used to evaluate the forecasts generated by the developed model against the actual observations.

To build an ARIMA model for the logged series, its sample autocorrelation function, ACF, (Fig. 3) and sample partial autocorrelation function, PACF, (Figure 4) are visually inspected.

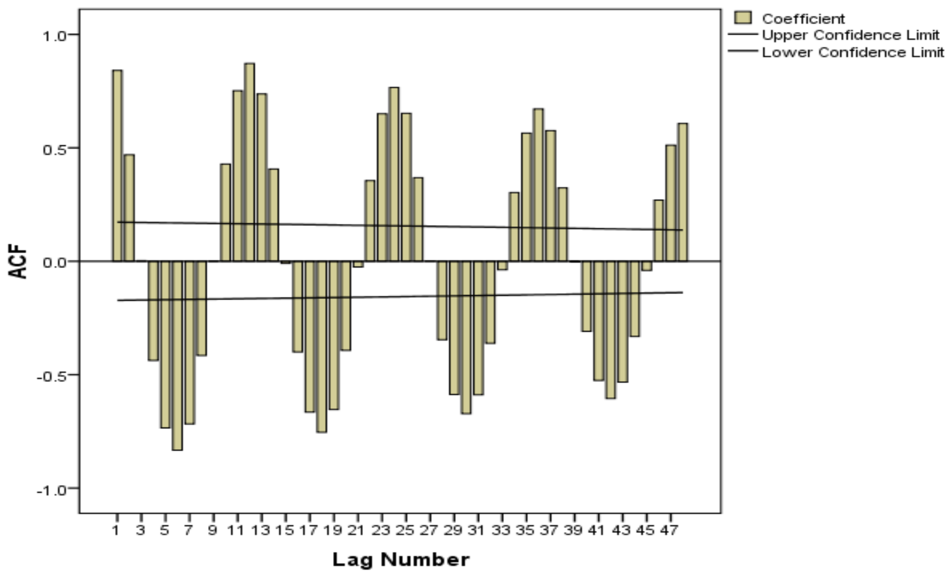



Fig. 3: The ACF of the Unadjusted logged Series

:

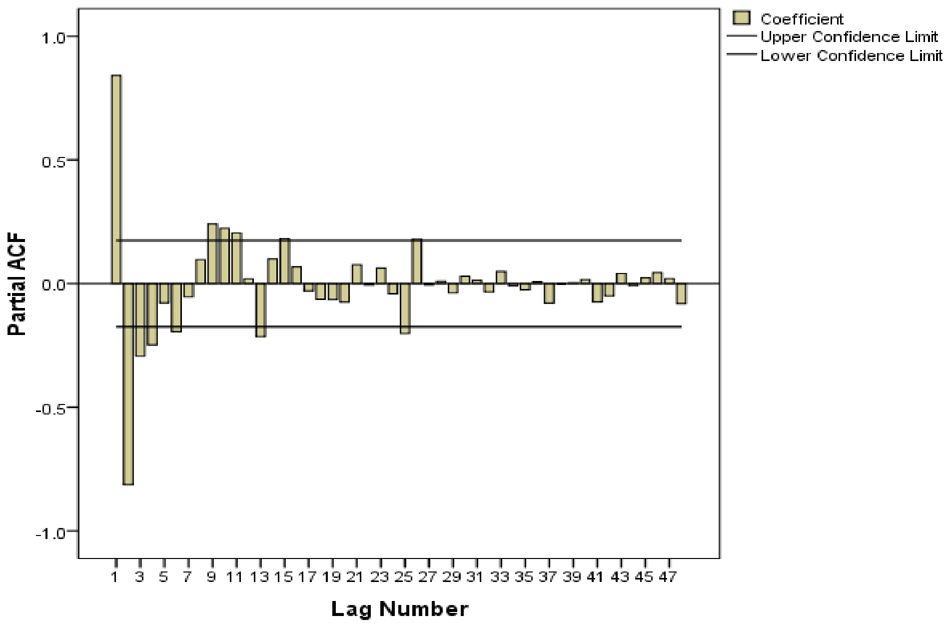
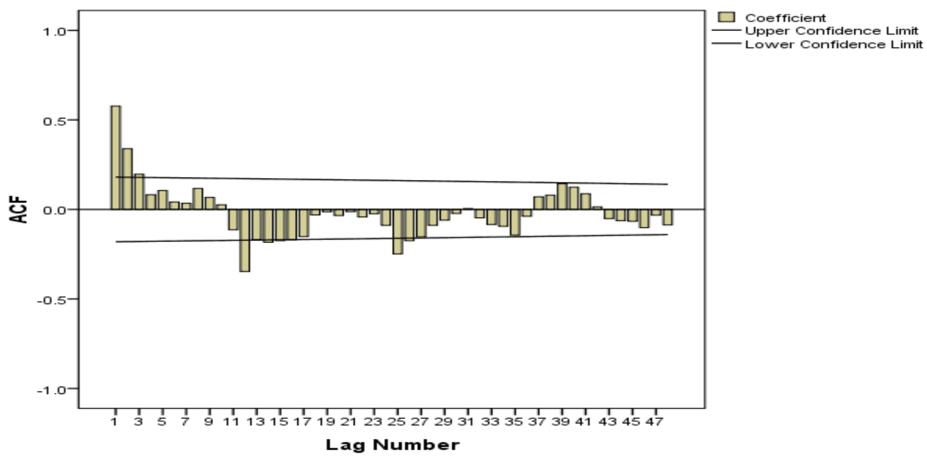



Fig. 4: The PACF of the Unadjusted logged Series

The ACF of the logged series is characterized by period of 12 and a failure to damp out, indicating that, the series is no stationary and the seasonal differencing is needed. The PACF also illustrates this nonstationarity by its ten large values, compared with its standard errors, found in lags 1, 2, 3, 4, 6, 9, 10, 11, 13, and 24. Since logged series does not seem to have a trend, any regular differencing is definitely inappropriate.

Denoting the original series by $X_t$, and the logged series by Ln $X_t$, a first seasonal difference of the logged series is $\nabla_{12}\text{Ln } X_t$ . The ACF of $\nabla_{12}\text{Ln } X_t$ is given in Fig.(5). It shows that no further seasonal differencing is required and that the series $\nabla_{12}\text{Ln } X_t$ is approximately stationary.

Fig. 5: The ACF of the Unadjusted logged Series with First Seasonal Difference




Since the ACF has values at lags 1, 2, and 3 depicting a rapid decay, and large peak at lag 12, an initial model

$$(1 - \phi_1 B)\ \nabla_{12} Ln\ X_t = (1 - \theta_1 B^{12})\ a_t \qquad (10)$$

is suggested. Since there is no steady trend in the logged series, no constant is included in the model.

Inspection of the PACF of $\nabla_{12} Ln\ X_t$ ( Fig. (6) also reveals that the series is stationary. The high spike at lag one suggests an autoregressive operator of order one, and the high spikes at lags 12, 13, and 24 suggest the inclusion of seasonal moving average operator of order one. Since the first seasonal difference of the logged series is stationary, no further seasonal difference is required.

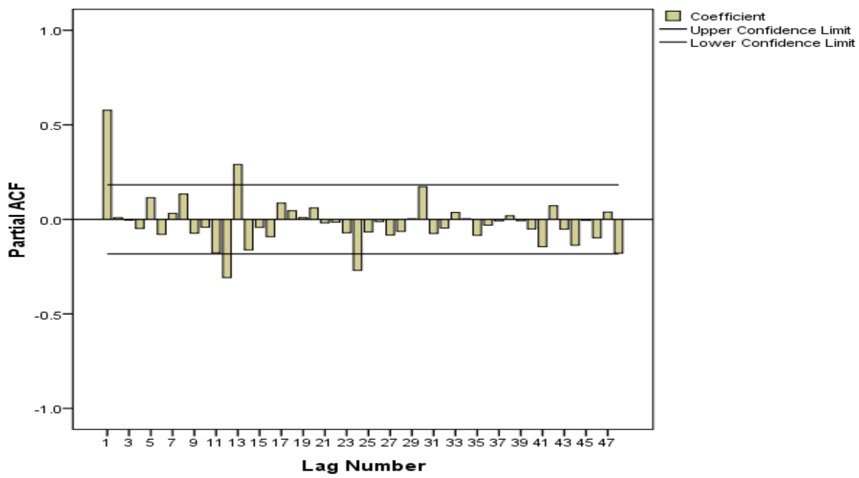



Fig. 6: The PACF of the Unadjusted logged Series with First Seasonal Difference

## Model Estimation

In the estimation phase, the parameters of the candidate models are estimated and their properties are examined in order to choose among them the best model. However, sometime the appropriate *ARIMA* model can be identified correctly after only a simple visual inspection of the ACF and the PACF of the time series in question.

The estimated *ARIMA* model for logged series with first seasonal difference for the period 1990-2000 and its corresponding statistics are given below. The standard errors are given in parentheses, and $Q^*(18)$ is the Box-Ljung statistics (computed for the first 18 lags) and is distributed as $\chi^2$ with 16 degrees of freedom. The sum of the squared residuals (*SSE*) from the estimated model is also reported in the estimation. The estimates of both the autoregressive and seasonal moving average parameters are found significant at 5% level of significance and their t-values are 10.020 and 6.138 respectively.

$$(1-0.697\,B)\,\nabla_{12} Ln\, X_t = (1-0.840\,B^{12})a_t$$

*(0.070)* *(0.137)*

$\sigma_a^2 = 0.0596$ $Q^*(16) = 17.996$ $SSE = 7.156$

**Model Diagnostic Checking**

Once the estimates of the model parameters are found, a diagnostic check is performed to test whether the estimated model is adequate. Inspection of the residual ACF (Fig. 7) of the estimated model reveals that there are no lags fall outside the bounds $\pm 1.96/\sqrt{n}$. In addition, the corresponding $Q^*$-statistic (17.996 based on first 18 lags) is less than the critical $\chi^2$ value ( 26.296 based on 16 degrees of freedom). The aforementioned results indicate that the model is adequate, and can, be used for forecasting purposes.

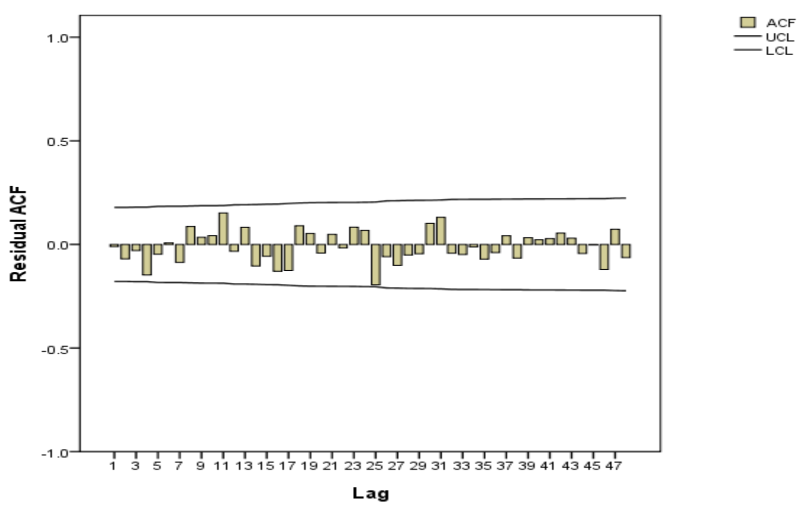



Fig. 7: Residual ACF of the Unadjusted logged Series with First Seasonal

**An ARIMA Model for the Seasonally Adjusted Data**

**Model Identification**

To construct the ARIMA model for the seasonally adjusted data, the seasonal variations are first removed from the logged series by applying the Census Method II. Since the seasonal variations are assumed to be captured by the seasonal adjustment procedure, no seasonal differencing will be applied to the resulting data prior to modeling.

From inspecting the ACF of the adjusted data (Fig. 8), it is found that significant spikes at lags 1, 2, 3, 8, 9, 10, and 11. The ACF values at lags 1, 2, and 3 exhibits a rapid decay and those at the lags 8, 9, 10, and 11 are considerably small.

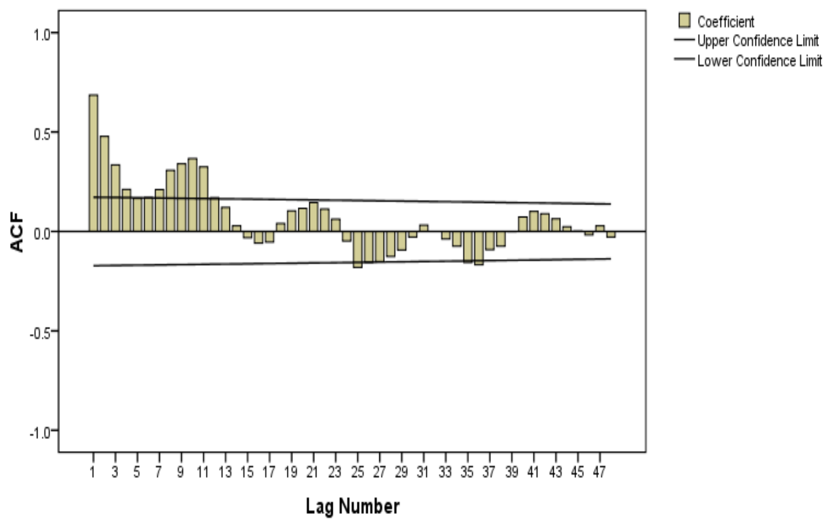



Fig. 8: The ACF of the Seasonally Adjusted logged Series

Inspection of the PACF (Fig. 9) reveals that it has significant spikes at lag 1. The pattern of the ACF for the first three lags and the high spike of the lag 1 in the PACF plot suggest the autoregressive operator of order one a tentative model. Denoting the logged series that is seasonally adjusted by $LnX_t^*$ , atentative model be in the form

$$(1 - \phi_1 B) LnX_t^* = a_t$$

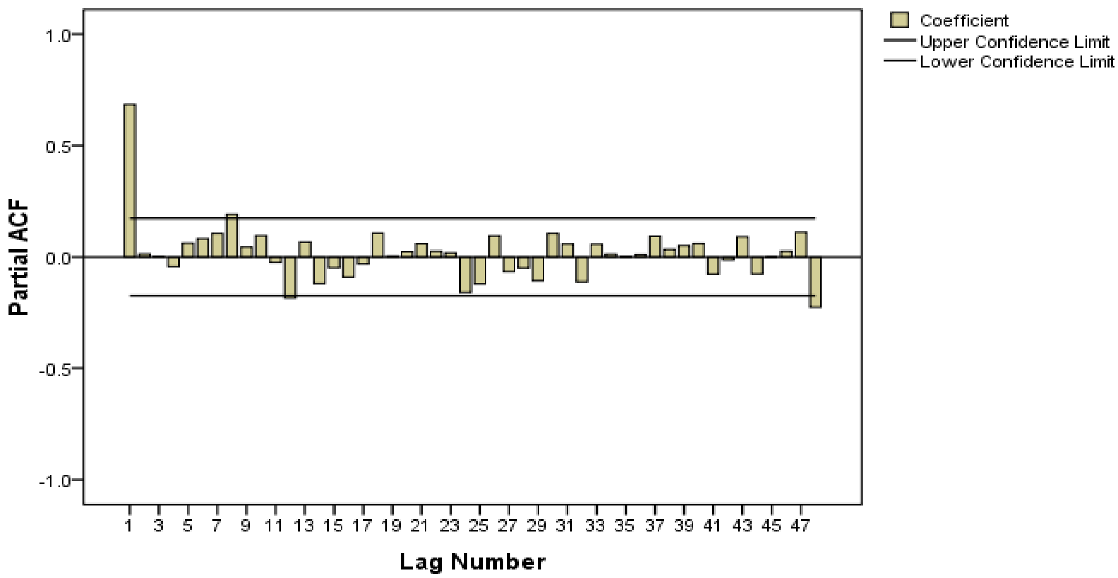



Fig. 9: The PACF of the Seasonally Adjusted logged Series

**Model Estimation**

The estimate of the autoregressive is found significant at 5 percent level of significance and its t-value is 728.059.

$(1 - 0.999B) LnX_t^* = a_t$ (0.001)

$\sigma_a^2 = 0.0659$ $Q^*(18) = 2.453$ $SSE = 9.053$

**Model Diagnostic Checking**

Examination of the residual ACF (Fig. 10) of the fitted model of seasonally adjusted series reveals that none is significant. Furthermore, the Box-Ljung statistic *Q*(18)* which has a value 2.453 for the first 18 lags, is far below the critical $\chi^2$ value (27.587 based on 17 degrees of freedom). These two results indicate that the fitted model is adequate and can, therefore, be used for forecasting purposes.

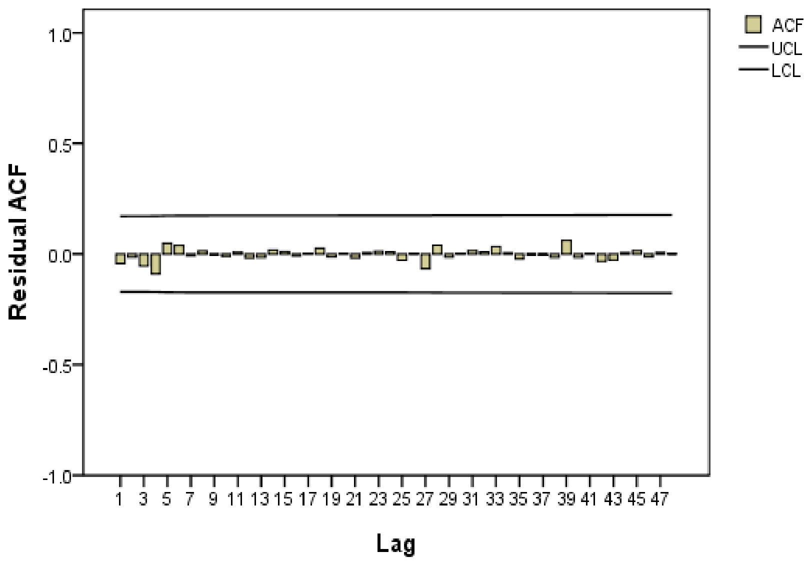



Fig. 10: Residual ACF of the Seasonally Adjusted logged

Evaluation of ARIMA Forecasts for Seasonally Adjusted and Unadjusted Series

The comparison of the forecasting performance of the fitted models for the adjusted and unadjusted series will be in terms of the residual variances, forecast error means, and mean absolute error (MAE) and root-mean square error (RMSE).

The residual variance from the fitted model for the unadjusted data is 0.0596, while it is *0.0659* for the adjusted data. A conclusion that the forecasts based on the unadjusted data will lead to more efficient predictions, therefore, be drawn.

Since the Census Method II seasonal Adjustment procedure equates the annual sums (Plosser, 1979), and therefore, the monthly averages of the seasonally adjusted and unadjusted series, forecasts in the form of averages of the 12 monthly predictions for each year through the forecasting interval are generated. After generating the forecasts for each specific year, the model are then updated by including the observations of that year in the estimation interval and the forecasts of next year are generated. This will enable making direct comparison between the forecasts based on the two forms of data.

The forecasts made by the models for both the adjusted and unadjusted data together with the actual values taken as monthly averages are given in table (1). Regarding the forecasting efficiency of the models as measured by the least absolute deviation of forecast from actual value for each year, it is found that the model from unadjusted data efficient in one years and the model from adjusted data efficient in five years.

Table 1: The Logged Series of Water Flow Data (Averages of Monthly Figures)

| Year | Actual | Prediction Unadjusted Series | Prediction Adjusted Series |
|---|---|---|---|
| 2001 | 7.52 | 7.64 | 7.83 |
| 2002 | 7.52 | 7.54 | 7.62 |
| 2003 | 7.49 | 7.51 | 7.35 |
| 2004 | 7.50 | 7.53 | 7.35 |
| 2005 | 7.42 | 7.53 | 7.38 |
| 2006 | 7.50 | 7.50 | 7.38 |

Table 2: Forecasting Model Coefficients (2001 – 2006) (The Unadjusted Data)

| Years | $\phi_1$ | $\theta_1$ | Adjusted Sum 0f Squares | Residual Variance |
|---|---|---|---|---|
| 2001 | 0.643 | 0.998 | 9.517 | 0.073 |
| 2002 | 0.545 | 0.838 | 13.129 | 0.091 |
| 2003 | 0.532 | 0.809 | 13.988 | 0.098 |
| 2004 | 0.523 | 0.755 | 14.705 | 0.089 |
| 2005 | 0.521 | 0.751 | 14.874 | 0.084 |
| 2006 | 0.537 | 0.820 | 15.867 | 0.084 |

A paired t-test for the actual monthly averages with their corresponding predicted values for the unadjusted and seasonally adjusted data resulted in t-values equal to -2.384 and 0.090, with probability values equal to 0.063 and 0.932 respectively. The same test for the values of the predictions of the unadjusted data with those of the adjusted, resulted in t-value equal to 0.903 with probability value equal to 0.408. The tabulated t-value for 5 degrees of freedom, at 0.05 level of significance, equal to 2.571.

Table 3: Forecasting Model Coefficient (2001 – 2006) (The Adjusted Data)

| Years | $\phi_1$ | Adjusted Sum 0f Squares | Residual Variance |
|---|---|---|---|
| 2001 | 0.999 | 13.495 | 0.090 |
| 2002 | 0.999 | 17.341 | 0.107 |
| 2003 | 0.999 | 17.431 | 0.100 |
| 2004 | 0.999 | 17.471 | 0.094 |
| 2005 | 0.999 | 17.567 | 0.089 |
| 2006 | 0.999 | 19.672 | 0.094 |

Therefore, it can be concluded that all of the differences between the matched pairs of the monthly averages are approximately equal to zero. The resulting coefficients together with the adjusted residual sum of squares and residual variance obtained upon the process of updating the parameter estimates for the years 2001 through 2006, are given in Table (2) and Table (3). A diagnostic check after each updating is performed to see whether the models degenerate over time and it is found that both models passed it.

The mean absolute error (MAE), root mean square error (RMSE) and the residual standard error (SE) for both forms of data are given in Table (4) and Table (5). Examination of these error statistics will help decide upon the efficiency of seasonal adjustment. Although both of the models passed the diagnostic check after each yearly updating , Table (2) and Table (3) reveal that the corresponding adjusted sum of squares and the residual variance are always much smaller than those of the adjusted data.

Table 4: Summary Statistics of Prediction Errors (The Unadjusted Data)

| Years | MAE | RMSE | SE |
|---|---|---|---|
| 2002 | 0.219 | 0.308 | 0.301 |

| | | | |
|---|---|---|---|
| 2003 | 0.218 | 0.305 | 0.299 |
| 2004 | 0.212 | 0.300 | 0.299 |
| 2005 | 0.206 | 0.293 | 0.291 |
| 2006 | 0.205 | 0.289 | 0.290 |

Table 5: Summary Statistics of Prediction Errors (The Adjusted Data)

| Years | MAE | RMSE | SE |
|---|---|---|---|
| 2001 | 0.242 | 0.704 | 0.300 |
| 2002 | 0.250 | 0.694 | 0.327 |
| 2003 | 0.238 | 0.669 | 0.317 |
| 2004 | 0.227 | 0.674 | 0.306 |
| 2005 | 0.218 | 0.627 | 0.298 |
| 2006 | 0.223 | 0.616 | 0.306 |

Using the error statistics of Table (4) and table (5); MAE, RMSE, and SE in comparing the prediction performance of the two models, it is found that, through the forecasting interval, the values of all these terms are lower for the unadjusted data compared with those of the adjusted data.

## CONCLUSIONS

This paper uses the Box-Jenkins iterative methods to develop forecasting models and present an empirical investigation of their forecasting performance for the seasonally adjusted and unadjusted forms of the water flow amount at Eldaim Station.

Regarding the efficiency of seasonal adjustment, the evidence presented in the analysis of water flow amount at Eldaim Station suggests that the use of seasonally adjusted data does not yield consistent improvement in forecasting accuracy relative to forecasts based on the unadjusted data. Indeed, the unadjusted data produce forecasts with lower mean absolute error, lower root mean square error, and lower residual variance. These finding suggest that there is some benefit from direct modelling of seasonality, rather than using the U. S. Bureau of the Census program to adjust the data prior to analysis. This can be attributed to three reasons; firstly, seasonal adjustment introduces distortions to the data series, secondly, no guarantee that the seasonal factors will be entirely removed, and thirdly, the use of alinear model for data obtained by nonlinear adjustment process.

For the water flow amount at Eldaim Station and on the basis of the results of forecast evaluation, it is reasonable to recommend that the Water Resources Department uses the unadjusted data when building univariate models that are to be used in forecasting or control.

## REFFERENCES

**Box, G. E. P. , and Jenkins, G. M**. (1976), "Time Series Analysis: Forecasting and Control" (rev. ed.), San Francisco: Holden-Day.

**Cleveland, W. P. And Tiao**, G. C. (1976). “Decomposition of Seasonal Time Series: A Model for the Census X-II Program", *Journal of the American Statistical Association*, Vol. 71, No. 355, PP. 581-587.

**Harvey, A. C**. (1993). "Time Series Models", 2nd ed. NY, London, Harvest Wheat Sheaf.

**Mabert, Vincent A**.( 1974). " The Box-Jenkins Methodology of Time Series Analysis", USA: n. p.

**McKenzie, Sandra K**. (1984). " Concurrent Seasonal Adjustment With Census X-II", *Journal of Business &Economic Statistics*, Vol. 2, No. 3.

**Musa, A. G. Mohamed**, (1998). " The Forecasting Performance of ARIMA Models for Roseires DamWater Flow: Seasonally Adjusted vs Unadjusted Data", M. Sc. Thesis, Unpuplished.

**Pindyck, Roberts., and Rubinfeled, Daniel L**. (1981). "Econometric Models and Economic Forecasting", Tokyo: Mc Graw. Hill, Inc.

**Plosser, Charles I**. (1979). "Short-term Forecasting and Seasonal Adjustment", *Journal of the American Statistical Association*, Vol. 74, No. 365, PP. 15-23.

**Shiskin, Julius. et. al**. (1967). "The X-II Variant of the Census Method II Seasonal Adjustment Program", USA, Bureau of the Census, Technical Peper No. 15.

**Shummway, Robert H. and Stoffer, David S**. (2005). "Time Series Analysis and its

Applications", 2nd ed., Springer Science+Business Media, Inc.

**Wallis, Kenneth F**. (1974). "Seasonal Adjustment and Relations between Variables", *Journal of the Royal Statistical Society*, 69, pp. 18-31.

## تقييم الأداء التنبؤي لنماذج ARIMA في حالة إزالة الآثار الموسمية وعدم إزالتها

### الخلاصة

تناولت هذه الورقة إجراء تقييم الأداء التنبؤي لنماذج ARIMA أحادية المتغير في حالة إزالة الآثار الموسمية من البيانات وفي حالة عدم إزالتها. و لإزالة الآثار الموسمية استخدم برنامج مكتب الاحصاءت الأمريكي (Bureau of Census Method X-II). تم تحويل قراءات مناسيب المياه عند محطة الديم إلى سلسلة مستقرة بحساب اللوغريثم الطبيعي للبيانات قبل بناء نماذج ARIMA أحادية المتغير في حالة إزالة الآثار الموسمية من البيانات وفي حالة عدم إزالتها (بين يناير 1990 إلى ديسمبر 2000 ) ومن ثم إجراء المقارنة بين الأداء التنبؤي المبني على تحديث النماذج (بين 2001 إلى 2006). اختبار t الزوجي بين متوسطات القيم الحقيقية و التنبؤية لا يدل على وجود اختلاف في الأداء التنبؤي لنماذج ARIMA قبل وبعد إزالة الآثار الموسمية. كما أن مقارنة الإحصاءات الخاصة بأخطاء التنبوات الناتجة تدل على أن الأداء التنبؤي لنموذج ARIMA الذي يتم الحصول عليه قبل إزالة الآثار الموسمية أفضل من أداء النموذج الذي يتم الحصول عليه بعد إزالة الآثار الموسمية. استخدام البيانات قبل إزالة الآثار الموسمية عند بناء نماذج ARIMA للتنبؤ والتحكم لقراءات مناسيب المياه عند محطة الديم أفضل.